

# Markovian Description of Unbiased Polymer Translocation

Felipe Mondaini<sup>1,2</sup> and L. Moriconi<sup>1</sup>

<sup>1</sup>*Instituto de Física, Universidade Federal do Rio de Janeiro,*

*C.P. 68528, 21945-970, Rio de Janeiro, RJ, Brazil*

<sup>2</sup>*Centro Federal de Educação Tecnológica Celso Suckow da Fonseca,*

*UnED Angra dos Reis, Angra dos Reis, 23953-030, RJ, Brazil*

## Abstract

We perform, with the help of cloud computing resources, extensive Langevin simulations which provide compelling evidence in favor of a general markovian framework for unbiased polymer translocation. Our statistical analysis consists of careful evaluations of (i) two-point correlation functions of the translocation coordinate and (ii) the empirical probabilities of complete polymer translocation (taken as a function of the initial number of monomers on a given side of the membrane). We find good agreement with predictions derived from the Markov chain approach recently addressed in the literature by the present authors.

## I. INTRODUCTION

The phenomenon of polymer translocation through membrane pores has been the subject of a flurry of research activity in recent years [1], a fact related to its relevance to the understanding and development of important biotechnological processes, like DNA sequencing, gene therapy and cytoplasmic drug delivery in living cells [2–5]. Computer simulations have been playing a dominant role in most of the expressive literature of polymer translocation, which may be, roughly, classified into the topical issues of i) translocation driven by chemical potential gradients [3, 6–10], ii) translocation driven by external forces [10–13], and iii) unbiased translocation [9, 10, 14, 15]. One is usually interested to compute (and, eventually, to model) the scaling exponents of the power laws which are found to relate the average translocation time  $\tau$  to the polymer size  $N$ .

Unbiased translocation, where the diffusion of a polymer through a membrane pore occurs uniquely as the consequence of thermal fluctuations, is by far the most studied case and it is also our focus in this paper. Chuang, Kantor and Kardar (CKK) [16] have introduced a successful description of unbiased homopolymer translocation from the simple assumption that the polymer’s evolution does not take it far from its equilibrium states. The essential physical picture is that unbiased translocation is ultimately due to the diffusion of the polymer center of mass. The translocation time is, thus, assumed to scale precisely in the same way as the Rouse relaxation time [17], so that  $\tau \sim N^{1+2\nu}$ , where  $\nu$  is the well-known Flory exponent ( $\nu \simeq 0.588$  in three-dimensions).

Since the scaling behavior  $\tau \sim N^{1+2\nu}$  departs from the one of usual brownian diffusion,  $\tau \propto N^2$ , it has been eventually suggested [18, 19] that unbiased polymer translocation could not be modeled as a markovian process – in other words, memory effects should be taken into account as an essential ingredient in any kinematical description of polymer translocation. However, it is of crucial importance to note that particular values of the translocation exponent are not sufficient *per se* to rule out the markovian nature of polymer translocation. A Markov chain approach can be put forward which actually leads to the CKK translocation exponent and to a closed analytical expression for the probability of complete polymer translocation that stands in good agreement with results obtained from Langevin simulations [20].

Our central aim in this work is to subject the markovian hypothesis of unbiased poly-

mer translocation to a more stringent test. In essential words, we have considered statistical ensembles of three-dimensional polymer translocation realizations, taken from Langevin simulations, with sizes considerably larger than the ones presently found in the literature. Our statistical data sets have been produced within the Grid Initiatives for e-Science virtual communities in Europe and Latin America (GISELA), a cloud computing framework supported by several academic institutions.

This paper is organized as follows. In section II, we discuss the Langevin dynamics of translocating polymers, which are modeled as bead-spring chains of Lennard-Jones particles with the finite-extension nonlinear elastic (FENE) potential. As a starting point, our simulations are validated from evaluations of the mean translocation time, which is verified to be compatible with the CKK scaling prediction [16]. We, then, compute two-point correlators of the translocation coordinate and, from them, the memory time of monomer translocation events. The polymer diffusion through the membrane pore is noted to be free of anomalies for times larger than the memory time scale, clearly suggesting that unbiased polymer translocation is essentially Markovian. Once we have found, actually, that small monomer clusters are correlated during translocation, we take this information into account in section III to compare, with good agreement, the empirical probabilities of complete translocation for polymers of various sizes to the theoretical probabilities obtained from the Markov chain approach of Ref. [20]. In section IV, we summarize our results and point out directions of further research.

## II. LANGEVIN SIMULATIONS

In our Langevin simulations, the excluded-volume and van der Waals interactions between beads (monomers and membrane atoms) separated by a distance  $r$  are modeled through a repulsive Lennard-Jones (LJ) potential with cutoff at length  $2^{1/6}\sigma$ , where  $\sigma$  is the bead diameter:

$$U_{LJ}(r) = \begin{cases} 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6] + \epsilon, & \text{if } r \leq 2^{1/6}\sigma, \\ 0, & \text{if } r > 2^{1/6}\sigma. \end{cases} \quad (1)$$

Besides the LJ potential, consecutive monomers are subject to the Finite-Extension Nonlinear Elastic (FENE) potential,

$$U_F(r) = -\frac{1}{2}kR_0^2 \ln[1 - (r/R_0)^2]. \quad (2)$$

From the above definition, it is clear that the FENE potential does not allow the distance between consecutive monomers to become larger than  $R_0$ .

We have studied polymers with sizes up to 300 monomers, which translocate through a pore created by the remotion of a single atom at the center of an 80 x 80 monoatomic square lattice membrane. Translocation is dynamically described by the following Langevin equations,

$$m \frac{d^2 \vec{r}_i}{dt^2} = - \sum_{j \neq i} \vec{\nabla}_{\vec{r}_i} [U_{LJ}(r_{ij}) + U_F(r_{ij})] - \xi \frac{d \vec{r}_i}{dt} + \vec{F}_i(t) , \quad (3)$$

where  $r_{ij} = |\vec{r}_i - \vec{r}_j|$ ,  $\xi$  is the dissipative constant and  $\vec{F}_i(t)$  is a gaussian stochastic force which acts on the monomer with label  $i$ , completely defined from the expectation values

$$\begin{aligned} \langle \vec{F}_i(t) \rangle &= 0 , \\ \langle [\hat{n} \cdot \vec{F}_i(t)] [\hat{n}' \cdot \vec{F}_j(t')] \rangle &= 2 \hat{n} \cdot \hat{n}' k_B T \xi \delta_{ij} \delta(t - t') . \end{aligned} \quad (4)$$

Above,  $\hat{n}$  and  $\hat{n}'$  are arbitrary unit vectors, and  $k_B$  and  $T$  are the Boltzmann constant and the temperature, respectively. By means of a suitable regularization of the stochastic force, we have implemented a fourth-order Runge-Kutta scheme for the numerical simulation of the Langevin Equations (3). Our simulation parameters are:  $\epsilon = 1.0$ ,  $\sigma = 1.0$  ( $\sigma$  is also identified with the membrane lattice parameter),  $\xi = 0.7$ ,  $k = 7\epsilon/\sigma^2$ ,  $R_0 = 2\sigma$ ,  $k_B T = 1.2\epsilon$ . The simulation time step is taken to be  $3 \times 10^{-3} t_{LJ}$ , where  $t_{LJ} \equiv \sqrt{m\sigma^2/\epsilon}$  is the usual Lenard-Jones time scale. In the most general case, the initial configuration of the polymer has  $n$  monomers on the trans-side of the membrane and  $N - n$  on the cis-side. Translocation is allowed to start only after an initial stage of thermal equilibrium is reached for the whole polymer.

As a preliminary validation step, we have checked if the translocation samples produced from the Langevin simulations would lead, in fact, to mean translocation times that scale with the polymer size as expected on the grounds of the CKK phenomenological theory, i.e.,  $\tau \sim N^\alpha$  with  $\alpha = 1 + 2\nu$ . The polymers are initially prepared to be in thermal equilibrium with an equal number of monomers on both sides of the membrane. Mean translocation times have been obtained from averages taken over ensembles of 70 complete translocation processes for each given polymer size. The mean translocation time as a function of the polymer size is shown in Fig. 1, which in fact indicates a reasonable agreement with the CKK scaling exponent. Being confident on the Langevin simulational scheme, we are now

ready to move to the study of more subtle aspects of polymer translocation.

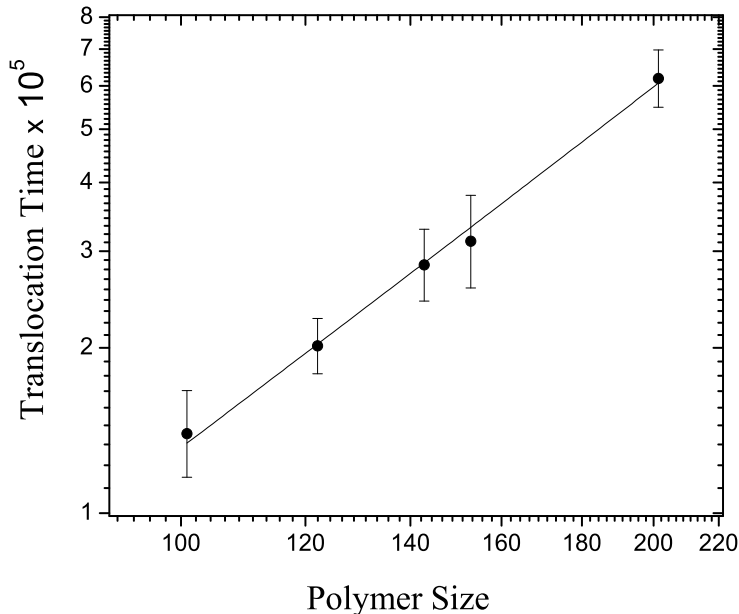


FIG. 1: The mean translocation time as a function of polymer size for unbiased translocation. We find the translocation exponent  $\alpha = 2.17 \pm 0.06$ , which is compatible with the CKK prediction,  $\alpha = 1 + 2\nu$ , with  $\nu = 0.588$  (numerical value of the three-dimensional Flory exponent).

### *Correlation Effects*

An important issue addressed in studies of polymer translocation refers to the role of correlations between monomer translocation events. As we show below, relying upon clear numerical evidence, such correlations are short-ranged in time, a fact that paves the way for a markovian description of unbiased translocation.

When a given dynamical system is claimed to evolve in time as a markovian stochastic process, it is of course tacitly assumed that the states of the system have been picked up in time intervals which are larger than some time scale associated to memory effects. In the polymer translocation context, it is convenient to define the dynamical state by the integer number  $s(t)$ , the “translocation coordinate”, that represents the number of monomers that have crossed the membrane to one of its sides, up to time  $t$  (which, by the way, is also treated as a discrete variable, since we measure it in units of the Lennard-Jones time scale). The

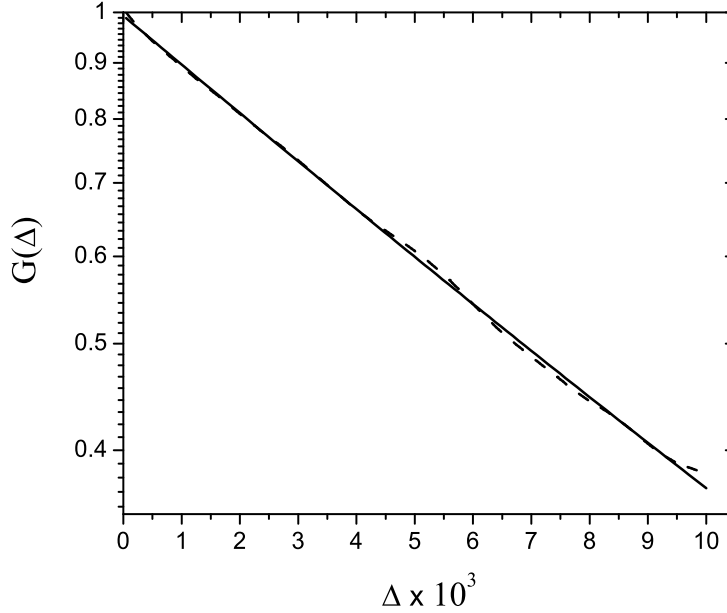


FIG. 2: The dashed line is the empirical evaluation of  $G(\Delta)$  as given by the truncated version of Eq. (5) with finite  $M$  and  $T$ . The solid line is the excellent fit provided by  $\exp(-\Delta/9300)$ .

memory time scale can be defined, in principle, from the decaying profile of the normalized two-point correlation function

$$G(\Delta) = \lim_{M, T \rightarrow \infty} \frac{\sum_{i=1}^M \sum_{t=0}^T s_i(t) s_i(t + \Delta)}{\sum_{i=1}^M \sum_{t=0}^T [s_i(t)]^2}, \quad (5)$$

where  $s_i(t)$  denotes the  $i$ -th sample taken from the ensemble of translocation coordinate time series. We have evaluated the right-hand side of (5) for polymers composed of  $N = 300$  monomers, in an ensemble of  $M = 1000$  samples, with time bound  $T = 1.5 \times 10^4$ . The initial polymer configurations are in thermal equilibrium with 150 monomers on each side of the membrane. As it is depicted in Fig. 2, the two-point correlation function (5) follows, to very good approximation, the simple exponential law  $G(\Delta) = \exp(-c\Delta)$ , with decaying time parameter  $1/c = 9.3 \times 10^3$ . In concrete terms, this result means that after around  $9.3 \times 10^3$  Langevin iterations, translocation loses memory of the past states. It is interesting, having in mind Markov modeling, to find the typical size of correlated monomer clusters that translocate within the memory time scale. As we discuss below, this can be achieved from an analysis of the diffusive behavior of polymer translocation.

### *The normal diffusive regime of unbiased translocation*

Still considering polymers of size  $N = 300$  in three dimensions, the variance of the translocation coordinate, that is,  $\langle [s(t)]^2 \rangle$ , is plotted in Fig. 3. Averages are now taken over ensembles of  $10^4$  samples. It is clear from that picture that for times larger than the memory time scale,  $\langle [s(t)]^2 \rangle$  is essentially a linear function of time or, in equivalent words, monomers diffuse in a normal way through the membrane pore.

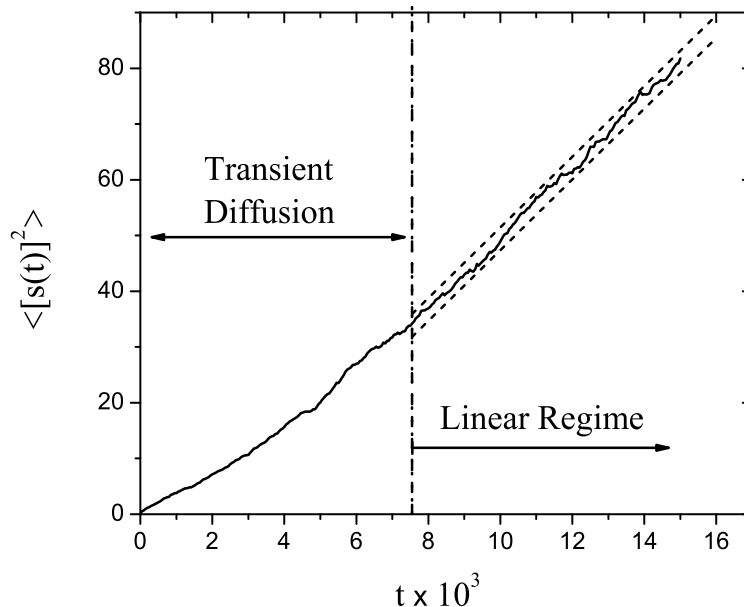


FIG. 3: Plot of the variance of the translocation coordinate as function of time for  $N=300$ . A linear regime effectively holds for  $t \gtrsim 8 \times 10^3$ , after an initial period of nonlinear transient behavior. Statistical averaging is performed in an ensemble of  $10^4$  independent polymer translocation realizations.

It is important to emphasize that our data provides a strong objection to the previously predicted anomalous scaling  $\langle [s(t)]^2 \rangle \sim t^{\frac{2}{1+2\nu}}$  [16], which must be viewed now as a misleading result derived within the CKK theory of the mean translocation time. Fig. 3 also indicates that around the memory time scale  $t = 9.3 \times 10^3$  we have  $\sqrt{\langle [s(t)]^2 \rangle} \simeq 6.5$ , which means that the correlated monomer clusters are composed of approximately six monomers (we choose to round the size of the monomer clusters to six and not to seven, once we observe that the crossover to the linear regime in Fig. 3 takes place a little before the memory time scale

set by the two-point correlation function (5)). The size of such correlated monomer clusters is a crucial ingredient in the Markov chain approach to polymer translocation: if the original polymer is replaced by a “monomer-clustered” polymer (whose size is the original size divided by the size of correlated monomer clusters), then the translocation of monomer clusters, rather than individual monomers, is assumed to generate a truly markovian stochastic process.

### III. EVIDENCE OF MARKOVIAN BEHAVIOR

The fundamental hypothesis of the Markov chain approach to unbiased translocation is that monomers translocate in an uncorrelated way with probabilities  $p_n$  and  $q_n$  for cis→trans and trans→cis transitions, respectively (recall that  $n$ , as defined in Sec. II, is the number of monomers on the trans-side of the membrane). Following Ref. [20], one puts forward the transition probabilities

$$\begin{aligned} p_n &= \frac{c}{(N - n)^{\delta+2\nu-1}} , \\ q_n &= \frac{c}{n^{\delta+2\nu-1}} , \end{aligned} \tag{6}$$

where  $0 < c < 1$  is an arbitrary constant and  $\delta$  is a scaling exponent associated to finite-size corrections to the CKK scaling relations (it would follow, for instance, that  $\tau \sim N^{\delta+2\nu}$  for fixed  $\delta$  and large enough  $N$ ). It is also possible, from the definitions (6) and using exact results for general Markov chains [21], to find the probability  $P(N, n)$  of complete trans→cis polymer translocation,

$$P(N, n) = \frac{1 + \sum_{i=1}^{n-1} \prod_{j=1}^i \frac{q_j}{p_j}}{1 + \sum_{i=1}^{N-1} \prod_{j=1}^i \frac{q_j}{p_j}} . \tag{7}$$

We have tested the analytical prediction given by Eq. (7) for polymers of various sizes. As it is shown in Fig. 4, the comparison between the empirical and analytical probabilities is very satisfactory. The empirical probabilities have been evaluated from ensembles of 70 complete polymer translocation realizations. To understand Fig. 4, note that the parameter  $N$  to be substituted in (6) is not the original polymer size anymore. In (6),  $N$  is now taken as the effective size of the monomer-clustered polymer (in our particular case, as suggested by the results of Sec. II, it is just the original polymer size divided by 6). Therefore, for a given value  $n$  of the initial number of trans-monomers, a solid circle is plotted in Fig. 4,



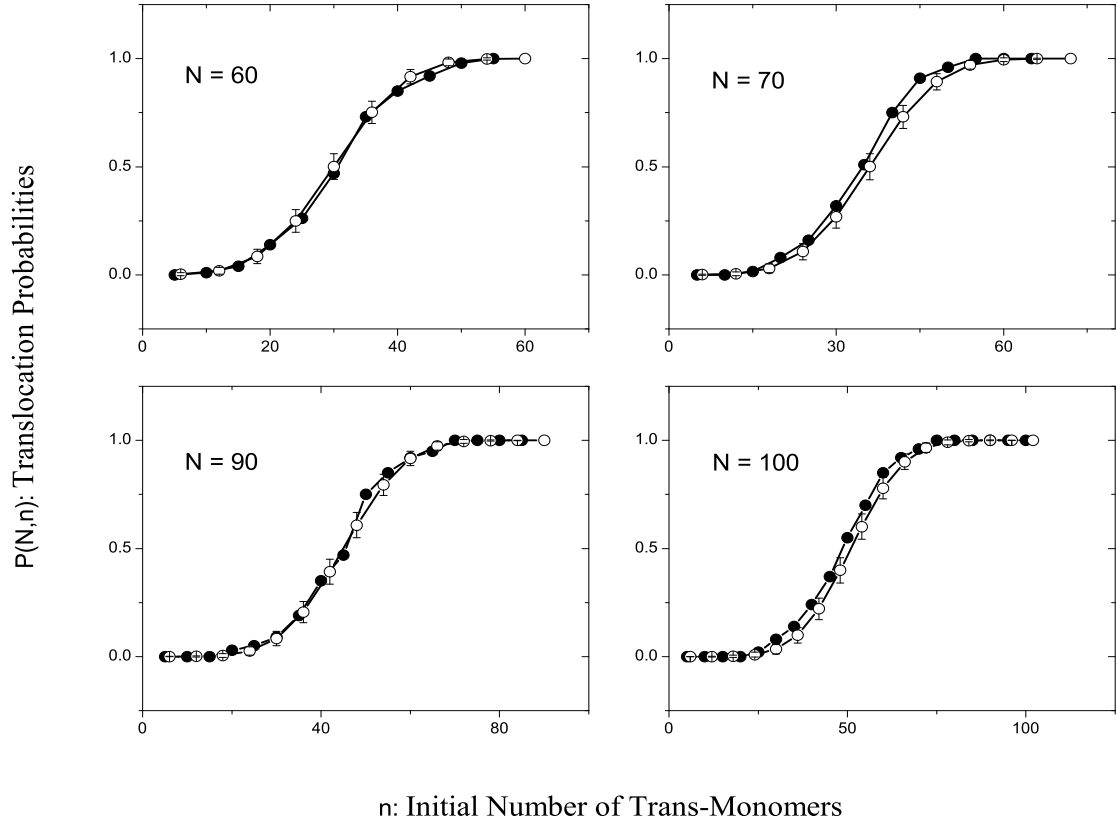


FIG. 4: The empirical (open circles) and analytical (solid circles) probabilities (both denoted here by  $P(N, n)$ ) of  $\text{cis} \rightarrow \text{trans}$  complete translocations are compared. The variable  $n$  stands for the initial number of monomers in the trans-side of the membrane. The analytical probabilities are computed from (6) and (7) with  $\delta = 0.88$  and  $\nu = 0.588$  for  $N = 60, 70, 90, 100$ . The theoretical error bars follow from elementary statistical considerations and are given by  $\sqrt{P(N, n)(1 - P(N, n))/70}$ .

with coordinates  $(n, P(N/6, n/6))$  where  $P(N/6, n/6)$  is the probability of complete polymer translocation evaluated within the Markov chain approach for a polymer which contains  $N/6$  monomers ( $N/6$  is conventionally rounded, if necessary to the smallest integer greater than  $N/6$ ).

## IV. CONCLUSIONS

We have provided consistent statistical data which essentially settles down the issue on whether unbiased polymer translocation is markovian or not – it turns out that unbiased polymer translocation can be very confidently described as a markovian stochastic process. This conclusion is supported from three clear pieces of evidence: (i) the two-point correlation function (5) indicates that individual monomer translocation events are correlated within a finite memory time scale which is much smaller than the mean complete translocation time; (ii) the variance of the translocation coordinate depends linearly on time for times which are larger than the memory time scale; and (iii) empirical probabilities of complete translocation finely match the analytical ones predicted from the Markov chain approach of Ref. [20] (it is worth of mentioning that (ii) can be also derived within the same formalism).

An interesting point, motivated by our results, is whether the original approach of polymer translocation addressed by Muthukumar [22], which is also markovian (but failed to give the correct expression for the mean translocation time), can be somehow improved taking into account the present findings. In concrete terms, it is not difficult to define a specific free-energy profile for the translocating polymer that would lead to the individual monomer translocation probabilities (6). A challenging problem which we are currently investigating is whether such effective free-energy profile is just an artifact that can be used to reproduce the stochastic evolution of the translocating polymer, or is actually the physical thermodynamical potential derived from standard equilibrium statistical mechanics considerations.

As a final remark, we call attention to recent works on driven translocation, where memory effects are found to be relevant, due to the excitation of collective modes along the polymer chain [11, 13]. However, as a problem deserved for further research, we note that if the forces which pull the polymer are not strong enough, it is likely that translocation can still be described with the help of Markov chain equations, where the transition probabilities (6) are just replaced by alternative ones.

This work has been partially supported by CNPq and FAPERJ.

---

[1] M. Muthukumar, *Annu. Rev. Biophys.* **36**, 435 (2007).

- [2] M. A. Zanta, P. Belguise-Valladier, and J.P. Behr, Proc. Natl. Acad. Sci. USA. **96**, 91 (1999).
- [3] J.J. Kasianowicz, E. Brandin, D. Branton, and D.W. Deamer, Proc. Natl. Acad. Sci. U S A. **93**, 13770-3 (1996).
- [4] D. Fologea, M. Gershow, B. Ledden, D.S. McNabb, J.A. Golovchenko, J. Li, Nano Lett. **5**, 1905 (2005).
- [5] K. Luo, T. Ala-Nissila, S.-C. Ying, and A. Bhattacharya, Phys. Rev. Lett. **100**, 058101 (2008).
- [6] A. Yu. Grosberg, S. Nechaev, M. Tamm and O. Vasilyev, Phys. Rev. Lett. **96**, 228105 (2006).
- [7] I. Huopaniemi, K. Luo, T.A.-Nissila, and S.-C. Ying, J. Chem. Phys. **125**, 124901 (2006).
- [8] K. Luo, T. Ala-Nissila, S.-C. Ying, and A. Bhattacharya, Phys. Rev. Lett. **99**, 148102 (2007).
- [9] D. Wei, W. Yang, X. Jin and Q. Liao, J. Chem. Phys. **126**, 204901 (2007).
- [10] K. Luo, S.T.T. Ollila, I. Huopaniemi, T. Ala-Nissila, P. Pomorski, M. Karttunen, S.-C. Ying, and A. Bhattacharya, Phys. Rev. E **78**, 050901(R) (2008).
- [11] D. Panja and G.T. Barkema, Biophys. J. **94**, 1630 (2008).
- [12] A. Bhattacharya, W.H. Morrison, K. Luo, T. Ala-Nissila, S.-C. Ying, A. Milchev, and K. Binder, Eur. Phys. J. E **29**, 423 (2009).
- [13] T. Ikonen, A. Bhattacharya, T. Ala-Nissila, and W. Sung, *Unifying Model of Driven Polymer Translocation*, <http://arxiv.org/abs/1111.4782>
- [14] D. Panja, G.T. Barkema and R.C. Ball, J. Phys. Condens. Matter **19**, 432202 (2007).
- [15] M.G. Gauthier and G.W. Slater, Phys. Rev. E **79**, 021802 (2009).
- [16] J. Chuang, Y. Kantor, and M. Kardar, Phys. Rev. E **65**, 011802 (2001).
- [17] P. E. Rouse, J. Chem. Phys. **21**, 1272 (1953).
- [18] D. Panja, J. Stat. Mech. P06011 (2010).
- [19] J. L. A. Dubbeldam, V. G. Rostiashvili, A. Milchev, and T. A. Vilgis, Phys. Rev. E **83**, 011802 (2011).
- [20] Felipe Mondaini and L. Moriconi, J. Chem. Phys. **135**, (2011).
- [21] N.G. van Kampen, *Stochastic Process in Physics and Chemistry*, Elsevier, Amsterdam (2007).
- [22] A. Muthukumar, J. Chem. Phys. **111**, 10371 (1999).